Applied Ocean Research 85 (2019) 45–52

Contents lists available at ScienceDirect

Applied Ocean Research

journal homepage: www.elsevier.com/locate/apor

A simplified hindcast method for the estimation of extreme storm surge events in semi-enclosed basins

D. Pasquali[a,*], M.F. Bruno[b], D. Celli[b], L. Damiani[b], M. Di Risio[a]

[a] University of L'Aquila, Department of Civil, Construction-Architectural and Environmental Engineering (DICEAA), Environmental and Maritime Hydraulic Laboratory (LIam), P.le Pontieri, 1, 67040 Monteluco di Roio, L'Aquila, Italy

[b] Technical University of Bari, Department of Civil, Environmental, Building Engineering and Chemistry (DICATECh), Coastal Engineering Laboratory, Area Universitaria di Valenzano S.P. Valenzano Casamassima, Km.3, 70010 Valenzano, Bari, Italy

ARTICLE INFO

Keywords:
Storm surge
Physics-based model
Sea level hindcast
ROMS
Coastal flooding risk
Climate change

ABSTRACT

Human presence, coastal erosion, and tourism activities are increasing the attention to coastal flooding risk. To perform risk assessments, long time series of observed or hindcast wave parameters and tide levels are then necessary. In some cases, only a few years of observation are available, so that observed extreme data are not always representative and reliable. A hindcast system aimed to reconstruct long time series of total tide levels may be of great help to perform robust extreme events analysis and then to protect human life, activities as well as to counteract coastal erosion by means of risk assessments. This work aims to propose a simplified method to hindcast storm surge levels time series in semi-enclosed basins with low computational costs. The method is an extension of a previous work of some of the authors and consists of a mixed approach in which the estimation of storm surge obtained by using the theory of linear dynamic system is corrected by using a statistical method. Both steps are characterized by low computational costs. Nevertheless, the results may be considered reliable enough also in view of the simplicity of the approach. The proposed method has been applied to the Manfredonia case study, a small village located in the Southern Adriatic Italian coast and often prone to coastal flooding events. The comparison of extreme events estimated on the basis of hindcast levels time series is satisfactorily similar to those estimated on the basis of observed tide series.

1. Introduction

The ongoing exploitation of coastal areas for touristic and socio-economic purposes, the related increase of human presence (e.g. [37]) and the increasing discussion on sea level rise due to climate change as well (e.g. [33]), have motivated more attention to coastal flooding. As a consequence, there is an increasing interest in the evaluation of coastal flooding risk. Furthermore, erosion phenomena are intimately related to coastal flooding events (e.g. [15]) and may cause stability problems to infrastructures near the coast, threatening touristic activities as well. Therefore, coastal managers and policy makers needs to have reliable tools to perform risk assessments in order to react against these factors.

Generally speaking, coastal risk is the mutual interaction between hazards and stakes [15], therefore it can be viewed as the probability that an event can happen and the evaluation of possible effects that it may induce. In literature, different definitions of risk concepts can be found. Among all, the risk may be defined as the product between three factors: hazard (the probability that an event with a defined return period could happen), exposure (the value of the elements contained in a defined area) and vulnerability (the capacity of a system to sustain hazardous events, or to react to them within the frame of resilience concept). Hence, the evaluation of the hazard is crucial to perform a reliable risk analysis. Indeed, all risk analysis tools need information about sea level conditions, intended as the superposition of astronomical tide (deterministic) and storm surge (stochastic), and sometimes information about wave height, in order to evaluate wave runup (e.g. [13,28]). However, these data are not always available (e.g. [25]) or, sometimes, the observed time series are not statistically representative for high return period. In other words, the observed time series are not long enough to be used to estimate storm surge with return periods up to several years. In these cases, it is common to use hindcast methods that allow long time series reconstruction, i.e. the topic this paper deals with. Usually, the state of the art of approaches for hindcast (especially at regional scale) is the use of hydrodynamic models forced by wind speed and pressure field from re-analysis data (e.g. [24]). However, these methods request high computational costs. The proposed method

* Corresponding author at: P.le Pontieri, 1, 67040 Monteluco di Roio, L'Aquila, Italy.
E-mail address: davide.pasquali@univaq.it (D. Pasquali).

https://doi.org/10.1016/j.apor.2019.01.031
Received 9 August 2018; Received in revised form 19 January 2019; Accepted 27 January 2019
Available online 05 February 2019
0141-1187/ © 2019 Elsevier Ltd. All rights reserved.

aims to be a low computational cost alternative to standard approaches by maintaining acceptable reliability.

As known, it is possible to reconstruct the harmonic component of the tide also with a short record; on the contrary, this does not apply to the estimation of the meteorological component of the tide (residual) due to its dependence on stochastic phenomena, i.e. wind field and pressure anomalies and the initial state of the system (e.g. [31,13]).

The main idea of this work is to hindcast the residual component of tide water levels by using the theory of linear dynamic systems (e.g. [17]) and by improving the reliability of the obtained results by means of statistical correction. In particular, this is an extension of the method proposed by Pasquali et al. [26] in which the same idea has been employed to forecast storm surge events in real-time, by using the response function to a unit wind stress of whatever the semi-enclosed basin. Their simplified model is able to forecast in real-time the surface elevation for wind time series by applying the discrete form of convolution integral based on wind speed forecast (given by whatever the meteorological center). However, as the meteorological component (residual) depends not only on the wind fields but also on the pressure anomalies, Pasquali et al. [26] proposed to correct the obtained time series by using a series of neural networks (one for each time delay in the forecast window).

The use of the same rationale is extended for hindcast purposes herein. It has to be stressed that hindcast is here devoted to have a long time series that can be used to estimate extreme events. Therefore, the timing of the reconstructed time series is not crucial as for real-time forecast, whilst the correct estimate of extreme levels is. More in details, the evaluation of the residual component due to the wind field relies on the same approach proposed by Pasquali et al. [26], by using hindcast data instead of forecast data. The model is extended by adding the inverse barometric effect and by proposing a different statistical correction. Based on the evaluation of the maximum error within a temporal window of given duration (e.g. [13]), hence not synchronous, the correction relies on the statistical selection of a calibration parameter aimed to provide a reliable estimate of extreme events. The proposed method is applied to a Point Of Interest (hereinafter referred to as POI) located in the Southern Adriatic Italian coast (Manfredonia). The obtained time series has been compared to measured data. A satisfactory accuracy in extreme events hindcast was observed.

The paper is structured as follows. Section 2 describes the method and its general features, Section 3 illustrates the features of the case study and available data, Section 4 discusses the obtained results. Concluding remarks close the paper.

2. The method for tide levels hindcast

The proposed method aims to hindcast sea levels using a physics-based approach, relying on the theory of linear dynamic systems coupled with a statistical correction. Since the main aim of the hindcast method is to build a statistically representative time series to perform robust extreme statistical analysis (e.g. for hazard evaluation), the main purpose of the proposed method is to correctly reproduce extreme values. Differently from Pasquali et al. [26], in this case it is not important to accurately reproduce synchronous events. Then, the method proposed by Pasquali et al. [26] has been extended and modified in view of these new requirements.

As known, tide levels are the superposition of an harmonic (deterministic) component, that may be estimated by means of standard harmonic analysis (e.g. [29,27,9]), and a stochastic one referred to as storm surge. The latter depends on meteorological (e.g. pressure anomalies, features of propagating perturbations) and morphological parameters (e.g. the presence of large continental shelves and their shape) that must be taken into account.

At first, the response function of a basin to a unit wind stress has to be computed by using a circulation model (this operation must be carried out once and for all). Later, under the hypothesis of linearity (reasonably acceptable in the case of semi-enclosed basins, e.g. [32]) it is possible to evaluate the response of the system to whatever the wind time series by using the proportionality and superposition principles. It has to be stressed that the wind time series, to be used within the theory of linear dynamic systems, has to be viewed as a series of impulses of finite duration.

More in details, the basin is discretized in a series of areas so that wind and pressure gradients can be caught.

The wind is considered constant over each area. For a more detailed description of the discretization, the reader may refer to Pasquali et al. [26].

Basically, wind impulses blowing in South-North (V_{ij}) and West-East (U_{ij}) directions, occurring between $t = (j-1)\Delta t$ and $t = j\Delta t$ ($j = 1,M$, being M the number of considered wind impulses) over the given area i ($i = 1,N$, being N the number of the areas) are considered. Then, the hindcast tide level time series due to the wind action at the POI can be expressed as the superposition of the contribution of each area. From a physical point of view, the obtained level takes into account only the wind field role in the storm surge generation process, so it cannot be considered reliable enough. However, there are other meteorological parameters (e.g. pressure anomalies, features of propagating perturbations) that influence the storm surge.

The residuals obtained by considering only the wind field role may be corrected by taking into account also the inverse barometric effects. Then, the equation able to hindcast tide level time series reads as follows:

$$\eta(t_k) = \sum_{i=1}^{N} \sum_{j=1}^{j\le M} U_{ij}^{(\tau)} F_i^U (t_k - j\Delta t + \Delta t) + V_{ij}^{(\tau)} F_i^V (t_k - j\Delta t + \Delta t) + C_p \Delta p(t_k) \quad (1)$$

where F_i^U and F_i^V are the response functions to a unit wind field that blows along the two Cartesian directions in area i, t_k ($= k \cdot \Delta t$; $k = 1, n$; n is the number of time steps) is the generic instant; $U_{ij}^{(\tau)}$ and $V_{ij}^{(\tau)}$ are the components of wind stress impulse along the West-East and the North-South axis respectively, i is the considered area, j is related to the number of the considered impulses occurring between $(j-1)\Delta t$ and $j\Delta t$, Δp is the pressure anomaly, C_p is a coefficient that takes into account the correlation between pressure anomalies and residual level. The coefficient C_p may be estimated by comparing synchronous observed residual levels and pressure values in the POI (e.g. [20]). It has to be stressed that the time shift between pressure anomalies and the related level fluctuations are not considered by Eq. (1).

The main aim of the proposed method is to achieve reliable estimates of extreme events. Then, an appropriate hindcast error needs to be defined first, e.g. by comparing the maximum measured ($\eta_{\max}^M$) and hindcast values ($\eta_{\max}^H$) occurring within a given temporal window. It has to be noted that within a temporal window, the two maximum values (measured and hindcast) may not be synchronous. Therefore, it is possible to define the maximum error (ϵ_{max}) as the ratio between $\eta_{\max}^M$ and $\eta_{\max}^H$ (e.g. [13]). Actually, the maximum error within a given temporal window can be viewed as a calibration coefficient (hereinafter referred to as C_{cal}) to be applied to hindcast series in order to improve the reliability of the estimate:

$$C_{\text{cal}} = \frac{\eta_{\max}^M}{\eta_{\max}^H}. \quad (2)$$

The maximum error expressed by Eq. (2) can be considered as a random variable if a series of temporal windows of a given duration is considered. De Girolamo et al. [13] proposed a statistical technique aimed to minimize the missed alarms of their real-time forecasting system, by selecting a quantile of the Empirical Cumulative Distribution Function (hereinafter referred to as ECDF) of the random variable C_{cal}. Herein, the aim is to hindcast extreme levels. Then, the quantile of the ECDF of the random variable C_{cal} is proposed to be selected by comparing the extreme analysis results based on the observation and on the corrected hindcast time series. It has to be stressed that the method can

be applied only if observations are available, at least for a few years. Indeed, the selection of the quantile of C_{cal} and the value of C_p can be performed on the basis of observed total levels and atmospheric pressure time series.

The application phase of the proposed method reads as follows: (i) estimation of the response function of the basin to a unit wind stress (to be carried out once and for all); (ii) estimation of the value of the parameter C_p; (iii) estimation of the residual levels by means of Eq. (1); (iv) selection of the quantile of the ECDF of the random variable C_{cal} based on the results of extreme values analysis on the hindcast and observed time series; (v) correction of the hindcast time series; (vi) hindcast of the total tide level as a superposition of harmonic and (corrected) residual time series.

In order to apply the method, some information are needed: (i) hindcast wind field and pressure anomalies within the considered basin, as provided by whatever the meteorological center; (ii) total levels time series measured at the POI to select the quantile of the ECDF of C_{cal}; (iii) atmospheric pressures measured at the POI to estimate C_p.

3. The case study of Manfredonia: site description and data

The proposed method has been applied to the case study of Manfredonia. It is a village located in the Southern Adriatic Italian coast (41.38°*N*, 15.55°*E*, Fig. 1). The Gulf with the same name, located on the South-Eastern part of the Gargano promontory, is a shallow area with a significant increase of bathymetric gradient occurring seaward beyond the 30 m isobath (e.g. [35]). The coastal areas on the Gulf are mainly sandy, the hinterland is characterized by wetlands with a great environmental relevance (e.g. [2,36]).

The areas of coastline in the Gulf are exposed to erosion processes and also prone to the flooding from Ofanto river (e.g. [12]). Moreover, there are also coastal flooding due to the presence of an extended area near the coast with elevation equal and, sometimes, lower than the sea level. When the wind comes from South or South-East (Sirocco wind), in particular during severe meteorological events, rivers flow is influenced by the downstream boundary condition and several areas are flooded. Due to these events, local residences, economic activities and accommodation facilities suffered large damages.

More in details, the sandy coast near Manfredonia belongs to the littoral cell extending between the harbor of Margherita di Savoia and the northernmost harbor of Manfredonia (Fig. 1). Starting from the '50s, the beaches in that area experienced a strong shoreline retreat caused by a massive anthropic action. The construction of the harbor of Margherita di Savoia (in 1952), and in particular of the Eastern Breakwater, interrupted the longshore sediment transport coming from Ofanto river located about 5 km southward. As a result, an extended sandy beach has formed East of the port, followed by a strong expansion of the town. In the western area, due to the lack of sediments, a significant shoreline retreat has been recorded. At first, coastal erosion affected the beaches just downdrift the port and then has spread northwards.

At present, the eroded coastline extends from Margherita di Savoia to the beach in front of Ippocampo tourist village (41.5°*N*, 15.91°*E*, Fig. 1); going towards the port of Manfredonia to the North, the shoreline can be considered substantially stable with a slight advance.

Despite standard protection structures (e.g. [16,7]) deployed to counteract the erosion over the past 40 years, erosion problems remains still unresolved. All the interventions have been conceived on a local scale, with the aim of solving localized erosive phenomena, without taking into consideration their effects in the entire littoral cell, causing a general worsening of the situation. At some locations, the protection structures underwent modifications, but also these interventions did not result in any significant beach advance.

Moreover, due to the presence of low-lying and depressed areas, coastal flooding occurs also during ordinary storms. Furthermore, sea storms are often associated with significant rainfall events and floods are worsened by river overflow. After these events, the area looks like a lagoon with a very long residence time and considerable damages to existing tourist and industrial sites. The coastal flooding hazard is particularly significant in the area close to the village of Ippocampo and the outlet of Carapelle river.

The proposed method is suitable to perform an hazard analysis of the case study. Nevertheless, measured time series are needed. As far as

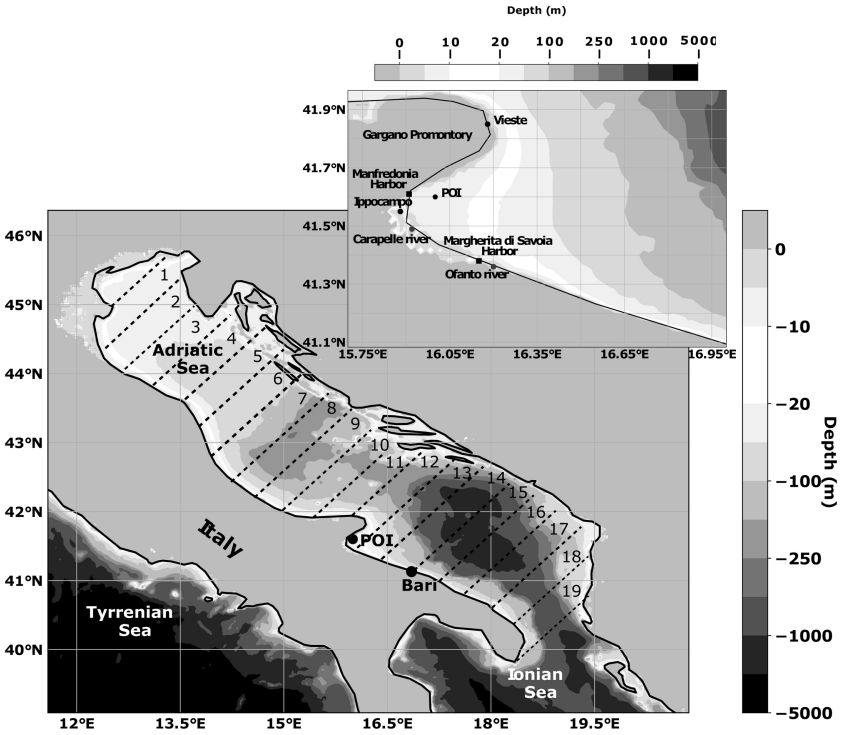

Fig. 1. Sketch of the discretization adopted for the Adriatic Sea used to compute the response functions. The wind has been applied along the main axis of the basin. Circles indicate the Point of interest (POI) and the locations mentioned in the paper. Dashed lines refer to the considered areas.

Table 1
Overview of the main information about the considered tidal stations.

Station	Period	Coordinates	Sampling time
Bari	01/01/1970 31/12/2012	41°08′24.74″N 16°51′57.72″E	1h
Vieste	01/01/1970 31/12/2012	41°53′17.10″N 16°10′37.24″E	1h
Manfredonia	01/01/2007 31/08/2015	41°37′28.39″N 15°54′51.27″E	0.25h

the tide levels are concerned, the measurements collected by means of a tidal gauge station belonging to the Apulia Region Meteomarine Network (also known as SIMOP, e.g. [3]) have been used. The network is operating since 2006, collecting wave, wind and tidal data along the Apulian coasts [11]. The tidal station is equipped with an ultrasonic gauge and a sensor to compensate for temperature changes. Identification of anomalous data, spikes and timing errors has been carried out by performing quality control checks (e.g. [19]), and spurious records have been removed from the time series.

The tidal station deployed at Manfredonia is not equipped with sensors to measure atmospheric pressures. Then, pressure data collected by means of two stations of the Italian National Tide Gauge Network installed near the POI are considered: Vieste and Bari (Fig. 1). Table 1 shows the details of the considered measurements.

The amplitude of tidal oscillations recorded in the Gulf of Manfredonia during the observation period (2007–2015) is of the order of tens of centimeters, consistent with sea level variations in the Mediterranean Sea. Moreover, as observed by past research works (e.g. [30,38,4]), interannual and long-term variability of the mean sea level has been observed both in the Mediterranean and the Adriatic Sea, strongly related to the North Atlantic Oscillation (NAO, e.g. [38]). Fig. 2 shows the mean sea level evaluated on the time series collected at Manfredonia, Bari and Vieste (see also Fig. 1) by applying a moving average filter with a time window of one-year duration. The results are consistent with previous research works (e.g. [30,38]). Then, all the analyses illustrated hereinafter are intended to be valid but the mean sea level, i.e. only residual levels are considered.

As far as the hindcast data in the Adriatic sea are concerned, reanalysis wind data and atmospheric pressure as provided by ERA-Iterim database (European Center for Medium-Range Weather Forecasts, e.g. [14]) have been used. It has to be underlined that the time resolution or ERA-Interim data is 6 h; this resolution does not allow to correctly reproduce very rapid meteorological events as long as correcting or calibrating the results against observations (e.g. [34]). The data were extracted from ERA-Interim database with a spatial resolution of 0.75°, whilst the response function may be estimated by performing numerical simulations with another spatial resolution. Therefore, the considered values of wind velocities and pressure acting at the center of each area may be obtained by means of linear interpolation (see next section).

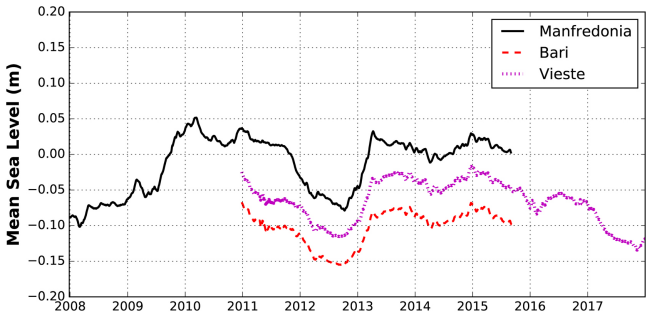

Fig. 2. Observed mean sea level evaluated on a moving average on time windows of one year duration.

4. Model application and discussion

This section aims to illustrate the application of the proposed method to the case study of the Manfredonia Gulf. The response functions of the system have been estimated through the numerical model Regional Ocean Modeling System (ROMS, e.g. [21,26]).

As known, storm surge events in the Adriatic Sea are due to Atlantic cyclones amplified by orographic and thermal effects (e.g. [6]). In particular, the configuration of the basin implies that the strongest storm surges are generated by pressure gradients and surface wind stress directed along the longitudinal axis of the basin (e.g. [6]). Furthermore, resonance phenomena may occur if meteorological anomalies force the natural modes of the basin (e.g. [8,5,39]). Then, also based on a preliminary sensitivity analysis, the basin has been discretized in 19 areas and the wind has been applied as the projection of the wind stress along the main axis of the basin (i.e. ≃324°N). Then, Eq. (1) is modified as follows:

$$\eta(t_k) = \sum_{i=1}^{N} \sum_{j=1}^{j\leq M} W_{ij}^{(\tau)} F_i^W (t_k - j\Delta t + \Delta t) + C_p \Delta p(t_k) \tag{3}$$

where F_i^W is the response function to a unit wind field stress that blows along the main axis of the basin in area i and $W_{ij}^{(\tau)}$ is the projection of the wind stress vector $\{U_{ij}^{(\tau)}, V_{ij}^{(\tau)}\}$ along the main axis of the basin. It has to be stressed that the number of areas has to be selected on a case by case basis. Indeed, it depends on the necessity to catch the wind and pressure gradients (e.g. a very coarse discretization does not allow to catch the variation of the atmospheric forcing), but also depends on the availability of the atmospheric data in the area, i.e. on the resolution of the selected Weather Forecast Centers (e.g. ECMWF, NOAA, etc.). In the case of the Adriatic Sea, Pasquali et al. [26] performed a sensitivity analysis to evaluate the most suitable number of areas showing that for more than 19 areas there were no significant variations in reliability.

More in details, to evaluate the response functions, each area has been forced using a unit wind stress impulse constant over the area, with a duration equal to 6 h (accordingly with the typical time resolution of General Circulation Models). It has to be stressed that due to the time resolution of the atmospheric data the model is not able to catch very rapid meteorological events (e.g. [34]). However, they are reliable enough to model typical perturbations in the Adriatic Sea that force the seiches of the considered semi-enclosed basin (e.g. [26]). The response function (with a temporal resolution equal to $\Delta\tau = 900$ s) could be extracted for a given POI as the free surface elevation time series. Fig. 3 shows the unit response functions evaluated for unit wind stress acting over two areas (Area 13 close to the POI, and Area 19 close to Otranto Strait, see Fig. 1).

The grid resolution is about 5500 m (175 × 185 computational points) and the bathymetric information has been defined based on 'ETOPO 1' database [1]. A total of 19 simulations have been carried out, where each simulation differs from the others for the area in which the wind stress impulse has been applied (constant over the area).

As underlined in Section 2, by using Eq. (3) it is possible to hindcast the residual level due to the action of wind stress and pressure anomalies.

In order to evaluate the correct values of the coefficient C_p, given by Eq. (3), for the considered POI, measured values (pressure and total tide level) are necessary. As underlined in Section 3, pressure measurements are not available in Manfredonia, while they are available near the POI, at Vieste and Bari (Fig. 1). By using the available pressure field data, a correlation between the (estimated) residual levels and the related measured pressure anomalies has been carried out. Fig. 4 shows the scatter diagrams and the estimation of the correlation coefficient C_p based on the measured pressure (P) and estimated residual at Vieste (upper panel) and Bari (lower panel).

As expected, despite their distances from the POI, the measured pressures are in good agreement between the two stations. This is

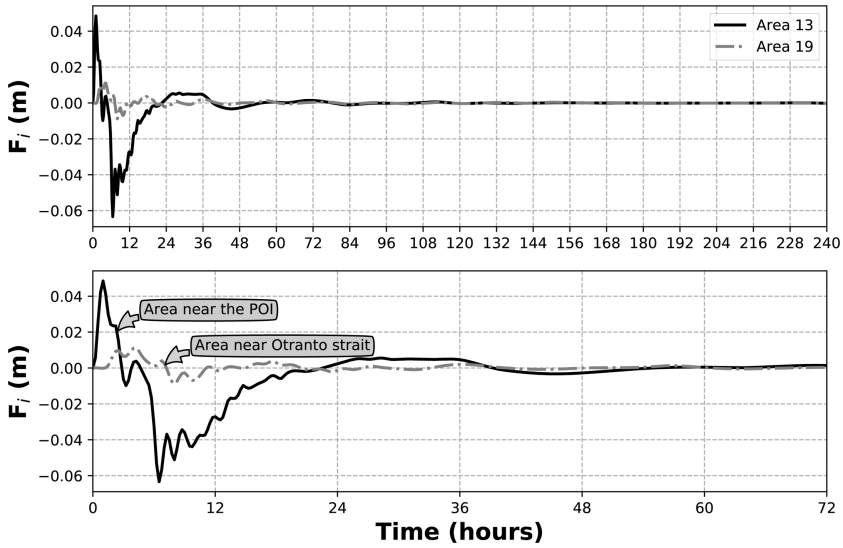

Fig. 3. Unit Response Functions of the basin evaluated by means of ROMS (upper panel) for areas 13 and 19 (see Fig. 1) and a detail of the first 72 h of the simulation (lower panel).

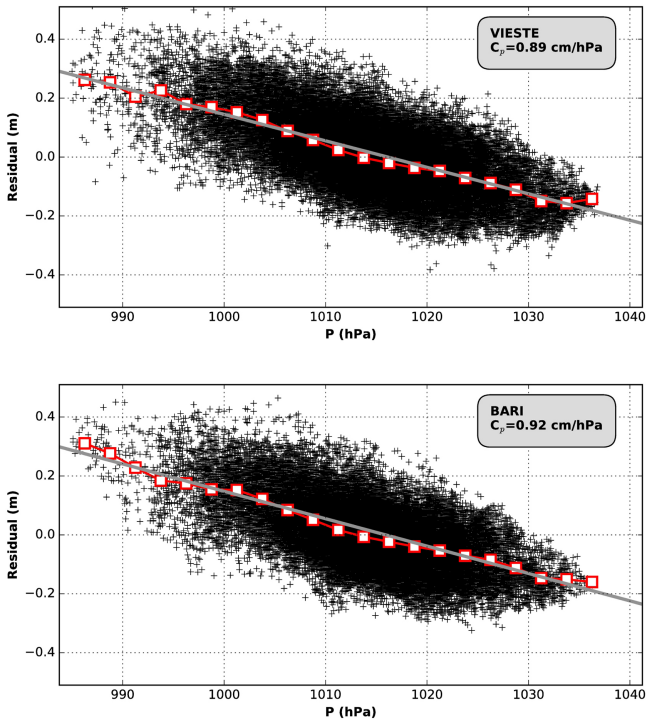

Fig. 4. Estimation of the coefficient C_p based on the measured pressure (P) and estimated residual at Vieste (upper panel) and Bari (lower panel). The grey line represents the best line fit of the average pressures evaluated on bins of 2.5 hPa (squares).

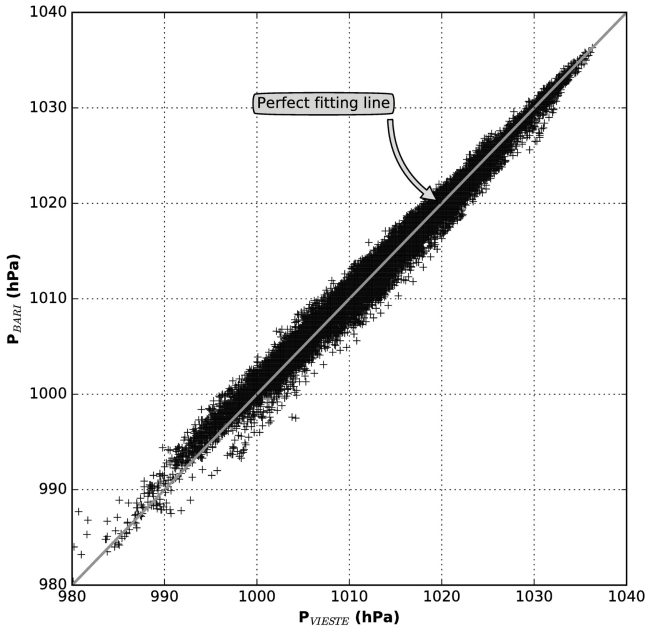

Fig. 5. Scatter plot of the synchronous measured pressure at Bari (P_{BARI}) and at Vieste (P_{VIESTE}). The grey line indicates the perfect fitting.

confirmed by the inspection of scatter plots of Fig. 5, where the synchronous measured pressures at Bari and Vieste are shown, and of Fig. 6 where synchronous pressure extracted from the ERA-Interim database at the center of area 12 (that is the area the POI belongs to, see Fig. 1) against the pressure observed at Bari (P_{BARI}, left panel) and at Vieste (P_{VIESTE}, right panel) are shown. It can be argued that from Vieste to Bari coast (about 150 km to the South) the pressure field does not change significantly. Therefore, hereinafter it is assumed that also in Manfredonia, located between Bari and Vieste, the pressures are almost similar to those measured in the two near mareographic stations. The value of C_p is then kept constant to 0.905 (the average value estimated for Bari and Vieste).

All data required to apply Eq. (3) are now known for the case at hand. Then, the hindcast of residual tide levels for a period of 39 years (from 1979 to 2017) has been evaluated. The roles of wind induced and pressure induced storm surge are strongly related to the specific features of the considered basin and to the features of perturbations. Hence, a general analysis of the importance of each term in Eq. (3) cannot be performed in a general sense. For the considered case study, the maximum contribution of the wind field may be roughly estimated to reach a value equal to a half of the maximum contribution of the pressure anomalies.

As underlined by Samayam et al. [34], ERA-Interim data could lead to an underestimation of the hindcast time series. This problem may be overcome by correcting or calibrating the data against observations. Therefore, as far as ERA-Interim data have been used, the resulting time series needs to be corrected in order to reach the requested reliability. A

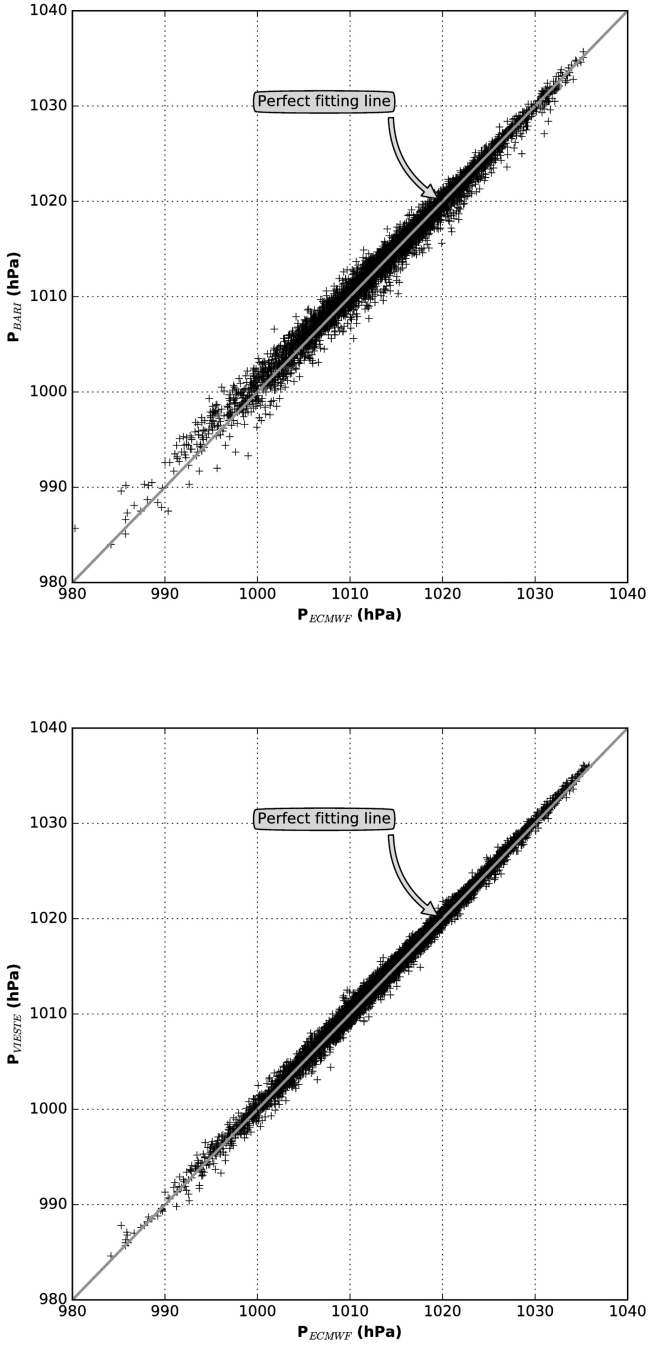

Fig. 6. Scatter plot of the synchronous pressure extracted from ERA-Interim at the center of area 12 against the pressure observed at Bari (P_{BARI}, upper panel) and at Vieste (P_{VIESTE}, lower panel). The grey lines indicate the perfect fitting.

Peak Over Threshold (POT) extraction was used to perform an extreme value analysis aimed to find the Generalized Pareto Distribution the data belong to (e.g. [10]). The threshold was selected by using the standard approach proposed by Coles et al. [10], i.e. basically, the highest value for which the shape and scale parameters remains almost unchanged.

Then, return levels were computed on the basis of extreme analysis performed on the observed values (Xr_M) and on the hindcast time series corrected with varying calibration coefficient (C_{cal} ranging from 1.0 up to 2.0, Xr_H). For the measured data, a threshold equal to 0.10 m with 162 values exceeding the threshold (i.e. about 4 extreme events per year) was identified. The estimated parameters of the GPD are equal to $\xi = 0.22$ (shape parameter) and $\sigma = 0.03$ (scale parameter).

Fig. 7 shows the variation of the ratio Xr_M/Xr_H as a function of the

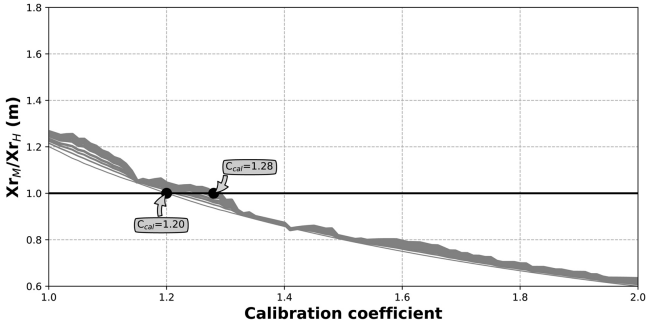

Fig. 7. Ratio of the return levels evaluated on the basis of the observed time series (xr_M) and of on the basis of the corrected hindcast time series (xr_H) as a function of the calibration coefficient C_{cal} used to correct the hindcast time series. The lines (with an increasing linewidth) refer to different return period (from 2 up to 100 years).

employed calibration coefficient (C_{cal}) applied to the hindcast time series. If a calibration coefficient equal to 1.24 ± 0.04 is considered, the ratio Xr_M/Xr_H is close to 1, i.e. the return levels evaluated on the basis of the corrected hindcast time series are almost equal to those evaluated on the basis of observed time series.

Fig. 8 shows the comparison for a series of four selected time windows of one month duration in the period 2007–2015 between observed residual level (thick grey lines), hindcast residual levels (dashed black lines) and corrected residual levels ($C_{cal} = 1.28$, black lines). From a qualitative point of view, it is possible to observe that the proposed method, even if characterized by a quite low computational cost, is able to reproduce observed storm surge events.

However, the method is aimed to get a reliable estimate of return levels. Then, Fig. 9 shows the Q-Q plots for $C_{cal} = 1$ (uncorrected hindcast series, left panel) and $C_{cal} = 1.28$ (right panel) in order to gain a deeper insight on the ECDF of the observed and hindcast extreme values.

To quantify the overall performances of the method in hindcasting extreme values, a series of statistical indicators has been used. The Root Mean Square Error (*RMSE*), the *Bias*, the Correlation Coefficient (*R*), the Index of Agreement (*d*) and the Nash Sutcliffe Efficiency Coefficient (*NSE*) have been calculated on the sample of the quantiles of the ECDF of the hindcast and observed extreme values (see Fig. 9). The first three parameters are commonly used in literature (e.g. [23,18,41]): *RMSE* is a measure of the spread of the data around the line of best fit, *Bias* indicates the spread around the mean and *R* is a measure of the correlation between predicted and observed values. Here other two indicators are used. The Index of Agreement (e.g. [40]) represents a measure of the error of the model. It can vary between 0 and 1; a value equal to 1 indicates a perfect accordance, 0 indicates no agreement. The Nash Sutcliffe Efficiency Coefficient is another measure used to evaluate the performances of a method to get reliable estimates, widely used to asses the goodness of fit (e.g. [22]). It can vary from $-\infty$ up to 1. A value equal to 1 corresponds to a perfect agreement between observed and predicted data. The definitions of the used indicators read as follows:

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i-O_i)^2} \quad \mathrm{BIAS}=\frac{1}{n}\sum_{i=1}^{n}P_i-O_i \quad \mathrm{R}=\frac{\sum_{i=1}^{n}(P_i-\bar{P})(O_i-\bar{O})}{\sqrt{\sum_{i=1}^{n}(P_i-\bar{P})^2(O_i-\bar{O})^2}}$$

$$\mathrm{d}=1-\frac{\sum_{i=1}^{n}(P_i-O_i)^2}{\sum_{i=1}^{n}(\mid P_i-\bar{O}\mid+\mid O_i-\bar{O}\mid)^2} \quad \mathrm{NSE}=1-\frac{\sum_{i=1}^{n}(O_i-P_i)^2}{\sum_{i=1}^{n}(O_i-\bar{O})^2} \tag{4}$$

where P_i ($i = 1, 2, \ldots n$) are the estimated values, O_i ($i = 1, 2, \ldots n$) are the observed ones, and n is the dimension of the sample.

Table 2 shows the results of the proposed model reliability by comparing the case in which the hindcast data have been corrected with a $C_{cal} = 1.28$ and the uncorrected case ($C_{cal} = 1$). The upper part of the table shows the values of the statistical indicators evaluated on the whole samples. It is possible to catch only a moderate increase in

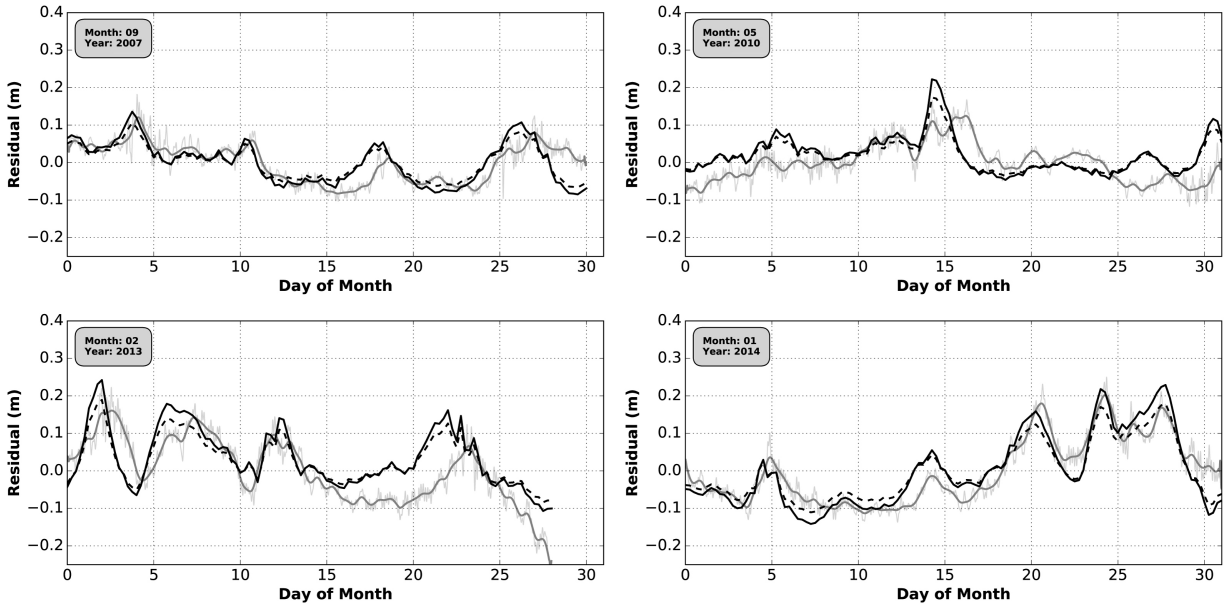

Fig. 8. Comparison of the observed (raw, thin grey lines; filtered, thick grey lines), hindcast (dashed black lines) and corrected hindcast (black lines) residual levels.

the reliability of the model using the calibration coefficient. Then, the indicators have been evaluated also considering only the quantiles greater than 0.15 m (lower part of the table) in order to better highlight the role of the statistical correction when low values are excluded from the analysis. The performance of the proposed method in hindcast extreme values is rather clear.

Table 3 shows the comparison between return levels evaluated on the basis of the measured levels and on the corrected hindcast, respectively. It could be observed that the return levels estimated for the observation time window (i.e. 2007–2015) are quite similar. Of course, this is expected as the hindcast time series was corrected in order to reproduce similar return levels. Nevertheless, it could be observed that also for the long duration corrected hindcast time series (i.e. 1979–2017), the estimated return levels do not change significantly. This is due to the occurrence of extreme events during the observation period. However, it should be underlined the improved statistical robustness, hence reliability, of the return levels estimation based on the long duration corrected hindcast time series with respect to those evaluated on the basis of the shorter observed time series.

Table 2
Statistical indicators of the efficiency and accuracy of the method with (C_{cal} = 1.28) and without (C_{cal} = 1.00) the application of the calibration coefficient.

Calibration	RMSE (m)	Bias (m)	R (–)	d (–)	NSE (–)
Quantiles over 0.10 m					
C_{cal} = 1.00	0.020	−0.012	0.987	0.952	0.853
C_{cal} = 1.28	0.010	0.007	0.991	0.990	0.960
Quantiles greater than 0.15 m					
C_{cal} = 1.00	0.030	−0.025	0.975	0.849	0.518
C_{cal} = 1.28	0.012	0.009	0.982	0.977	0.912

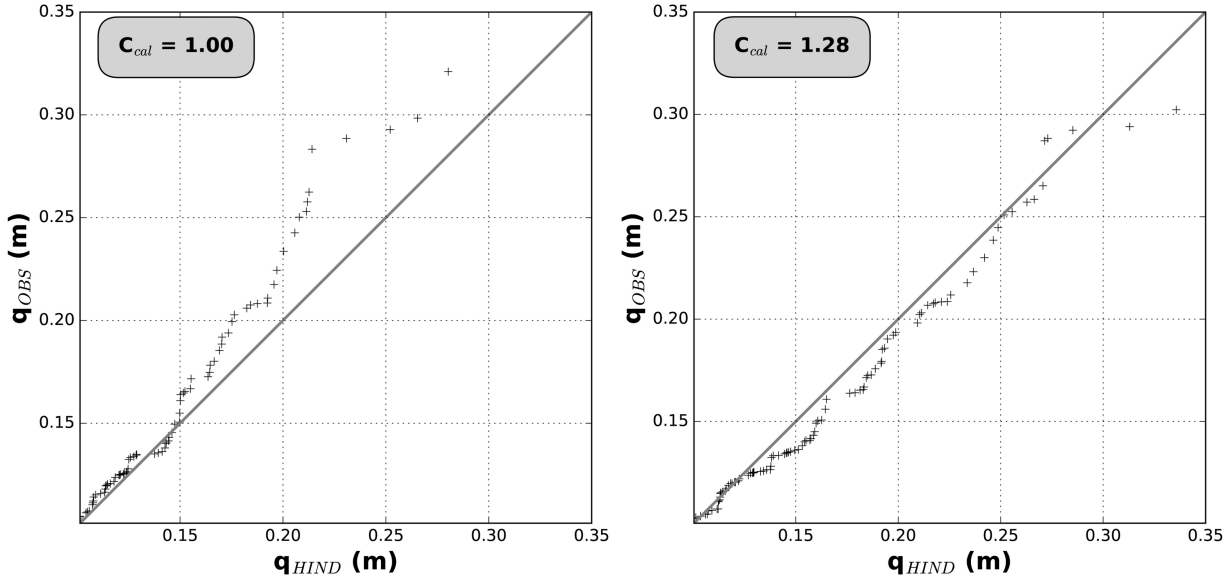

Fig. 9. Q-Q plots for C_{cal} = 1 (uncorrected hindcast series, left panel) and C_{cal} = 1.28 (right panel) of ECDF quantiles of the extracted extreme values.

Table 3
Return levels estimation based on GPD analysis for different return periods from 2 to 100 years.

Period	Data	2	5	10	20	50	100
2007–2015	Measures (m)	0.28	0.32	0.34	0.37	0.39	0.41
2007–2015	Hindcast (m)	0.30	0.34	0.36	0.38	0.40	0.41
1979–2017	Hindcast (m)	0.28	0.32	0.36	0.36	0.38	0.40

5. Concluding remarks

A simplified method for the estimation of return levels of storm surge when short observed time series are available is proposed.

The method relies on the use of a physics-based approach and a statistical correction. The former is based on the theory of linear dynamic systems applied to semi-enclosed basins. The latter consists of correcting the hindcast time series by means of a calibration coefficient aimed to reproduce the return levels evaluated on the basis of (short) observed time series.

The main idea is to compute (once and for all), by means of whatever the numerical model, the response functions at a selected Point of Interest of the considered basin to a unit wind stress. The response functions are then used to evaluate the response to whatever the wind time series, by computing a convolution integral within the frame of the linear dynamic systems theory (commonly accepted to describe the behavior of semi-enclosed basins). The method takes into account also the role of the inverse barometric effect. The obtained hindcast time series is then corrected by using a calibration coefficient selected in order to reproduce the return levels evaluated on the basis of (short) observed time series.

The proposed method may be applied if (i) tide levels and (ii) atmospheric pressures short time series (at least a few years) are available at (or close to) the point of interest and if (i) wind speed and (ii) atmospheric pressures hindcast time series are available within the semi-enclosed basin from whatever the meteorological center

The proposed method was applied to the case study of Manfredonia, an Italian coastal area located in the Italian Southern Adriatic coast. The practical application was intended to show the applicability of the method and its reliability, also in view of its quite low computational costs. Indeed, it was shown that the (corrected) hindcast time series can be used to achieve robust estimates of return levels of storm surge, useful, just for instance, for coastal risk assessments.

References

[1] C. Amante, B. Eakins, Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis (etopo1), NOAA, National Geophysical Data Center, Boulder, Colorado, USA, 2009.
[2] Apulia Region, Attività finalizzate alla redazione del piano regionale delle coste (prc) della regione puglia, Atti della Regione Puglia (2007) (in Italian).
[3] Apulia Region, Apulia Region Meteomarine Network, (2018) (accessed 06.05.18), http://www.adb.puglia.it/public/news.php?extend.282.6.
[4] M. Bruno, M. Molfetta, A. Petrillo, The influence of interannual variability of mean sea level in the adriatic sea on extreme values, J. Coast. Res. 70 (2014) 241–246.
[5] P. Canestrelli, M. Mandich, P. Pirazzoli, A. Tomasin, Venti, depressioni e sesse: perturbazioni delle maree a Venezia (1951–2000), Centro Previsioni e Segnalazioni Maree, 2001 (in Italian).
[6] G. Cecconi, P. Canestrelli, C. Corte, M. Di Donato, Climate record of storm surges in Venice, RIBAMOD-River Basin Modelling, Management and Flood Mitigation, Concerted Action, Proc. of the Final Wallingford Workshop, vol. 18287, European Commission, Luxembourg, EUR, 1999, pp. 149–156.
[7] D. Celli, D. Pasquali, P. De Girolamo, M. Di Risio, Effects of submerged berms on the stability of conventional rubble mound breakwaters, Coast. Eng. 136 (2018) 16–25.
[8] I. Cerovečki, M. Orlić, M.C. Hendershott, Adriatic seiche decay and energy loss to the mediterranean, Deep Sea Res. Part I: Oceanogr. Res. Papers 44 (1997) 2007–2029.
[9] D.L. Codiga, Unified Tidal Analysis and Prediction Using the UTide Matlab Functions. Technical Report 2011–01, Graduate School of Oceanography, University of Rhode Island Narragansett, RI, 2011.
[10] S. Coles, J. Bawa, L. Trenner, P. Dorazio, An Introduction to Statistical Modeling of Extreme Values Volume 208, Springer, 2001.
[11] L. Damiani, M.F. Bruno, M.G. Molfetta, B. Nobile, Coastal zone monitoring in apulia region: first analysis on meteomarine climate, Proceedings of 5th International Symposium on Environmental Hydraulics (ISEH 2007) Arizona (USA) (2007).
[12] L. Damiani, A. Petrillo, G. Ranieri, The Erosion Along The Apulian Coast Near the Ofanto River vol. 70, (2003), pp. 407–418.
[13] P. De Girolamo, M. Di Risio, G. Beltrami, G. Bellotti, D. Pasquali, The use of wave forecasts for maritime activities safety assessment, Appl. Ocean Res. 62 (2017) 18–26.
[14] D.P. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al., The era-interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc. 137 (2011) 553–597.
[15] M. Di Risio, A. Bruschi, I. Lisi, V. Pesarino, D. Pasquali, Comparative analysis of coastal flooding vulnerability and hazard assessment at national scale, J. Marine Sci. Eng. 5 (2017) 51.
[16] M. Di Risio, I. Lisi, G. Beltrami, P. De Girolamo, Physical modeling of the cross-shore short-term evolution of protected and unprotected beach nourishments, Ocean Eng. 37 (2010) 777–789.
[17] M. Di Risio, D. Pasquali, I. Lisi, A. Romano, M. Gabellini, P. De Girolamo, An analytical model for preliminary assessment of dredging-induced sediment plume of far-field evolution for spatial non homogeneous and time varying resuspension sources, Coast. Eng. 127 (2017) 106–118.
[18] T.H. Durrant, D.J. Greenslade, Spatial evaluations of access marine surface winds using scatterometer data, Aust. Meteor. Oceangr. J 62 (2012) 263–276.
[19] D. Evans, C. Conrad, F. Paul, Handbook of Automated Data Quality Control Checks and Procedures of the National Data Buoy Center, NOAA National Data Buoy Center Tech. Doc, 2003 03–02.
[20] O. Faggioni, G. Arena, N. Bencivenga, G. Bianco, R. Bozzano, G. Canepa, P. Lusiani, G. Nardone, G. Piangiamore, M. Soldani, The Newtonian approach in the meteorological tide waves forecasting: preliminary observations in the east ligurian harbours, Ann. Geophys. (2006) 49.
[21] D.B. Haidvogel, H. Arango, W.P. Budgell, B.D. Cornuelle, E. Curchitser, E. Di Lorenzo, K. Fennel, W.R. Geyer, A.J. Hermann, L. Lanerolle, et al., Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the regional ocean modeling system, J. Comput. Phys. 227 (2008) 3595–3624.
[22] R.H. McCuen, Z. Knight, A.G. Cutter, Evaluation of the nash-sutcliffe efficiency index, J. Hydrol. Eng. 11 (2006) 597–602.
[23] D.N. Moriasi, J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, T.L. Veith, Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Trans. ASABE 50 (2007) 885–900.
[24] S. Muis, M. Verlaan, H. Winsemius, J. Aerts, P. Ward, A global reanalysis of storm surges and extreme sea levels, Nat. Commun. (2016) 7.
[25] S. Muis, M. Verlaan, H. Winsemius, P. Ward, A first global-scale hindcast of extreme sea levels induced by extra-tropical storms, EGU General Assembly Conference Abstracts, vol. 17 (2015).
[26] D. Pasquali, M. Di Risio, P. De Girolamo, A simplified real time method to forecast semi-enclosed basins storm surge, Estuar. Coast. Shelf Sci. 165 (2015) 61–69.
[27] R. Pawlowicz, B. Beardsley, S. Lentz, Classical tidal harmonic analysis including error estimates in matlab using t_tide, Comput. Geosci. 28 (2002) 929–937.
[28] J. Perez, M. Menendez, I.J. Losada, Gow2: a global wave hindcast for coastal applications, Coast. Eng. 124 (2017) 1–11.
[29] D.T. Pugh, Tides, Surges and Mean Sea-Level (Reprinted With Corrections), John Wiley & Sons Ltd, 1996.
[30] F. Raicich, Recent evolution of sea-level extremes at trieste (northern adriatic), Cont. Shelf Res. 23 (2003) 225–235.
[31] W.E. Rogers, P.A. Wittmann, D.W. Wang, R.M. Clancy, Y.L. Hsu, Evaluations of global wave prediction at the fleet numerical meteorology and oceanography center, Weather Forecast. 20 (2005) 745–760.
[32] P.C. Roos, H.M. Schuttelaars, Influence of topography on tide propagation and amplification in semi-enclosed basins, Ocean Dyn. 61 (2011) 21–38.
[33] A. Salman, S. Lombardo, P. Doody, Living with coastal erosion in Europe: sediment and Space for Sustainability, EUCC (2004).
[34] S. Samayam, V. Laface, S.S. Annamalaisamy, F. Arena, S. Vallam, P.V. Gavrilovich, Assessment of reliability of extreme wave height prediction models, Nat. Hazards Earth Syst. Sci. 17 (2017) 409–421.
[35] F. Spagnoli, G. Bartholini, E. Dinelli, P. Giordano, Geochemistry and particle size of surface sediments of gulf of manfredonia (southern adriatic sea), Estuar. Coast. Shelf Sci. 80 (2008) 21–30.
[36] A. Specchiulli, F. Bignami, M. Marini, A. Fabbrocini, T. Scirocco, A. Campanelli, P. Penna, A. Santucci, R. D'Adamo, The role of forcing agents on biogeochemical variability along the southwestern adriatic coast: the gulf of manfredonia case study, Estuar. Coast. Shelf Sci. 183 (2016) 136–149.
[37] J.P. Syvitski, C.J. Vörösmarty, A.J. Kettner, P. Green, Impact of humans on the flux of terrestrial sediment to the global coastal ocean, Science 308 (2005) 376–380.
[38] M. Tsimplis, F. Calafat, M. Marcos, G. Jordà, D. Gomis, L. Fenoglio-Marc, M. Struglia, S. Josey, D. Chambers, The effect of the nao on sea level and on mass changes in the mediterranean sea, J. Geophys. Res.: Oceans 118 (2013) 944–952.
[39] I. Vilibić, The role of the fundamental seiche in the adriatic coastal floods, Cont. Shelf Res. 26 (2006) 206–216.
[40] C.J. Willmott, S.G. Ackleson, R.E. Davis, J.J. Feddema, K.M. Klink, D.R. Legates, J. O'donnell, C.M. Rowe, Statistics for the evaluation and comparison of models, J. Geophys. Res.: Oceans 90 (1985) 8995–9005.
[41] C.J. Willmott, S.M. Robeson, K. Matsuura, A refined index of model performance, Int. J. Climatol. 32 (2012) 2088–2094.